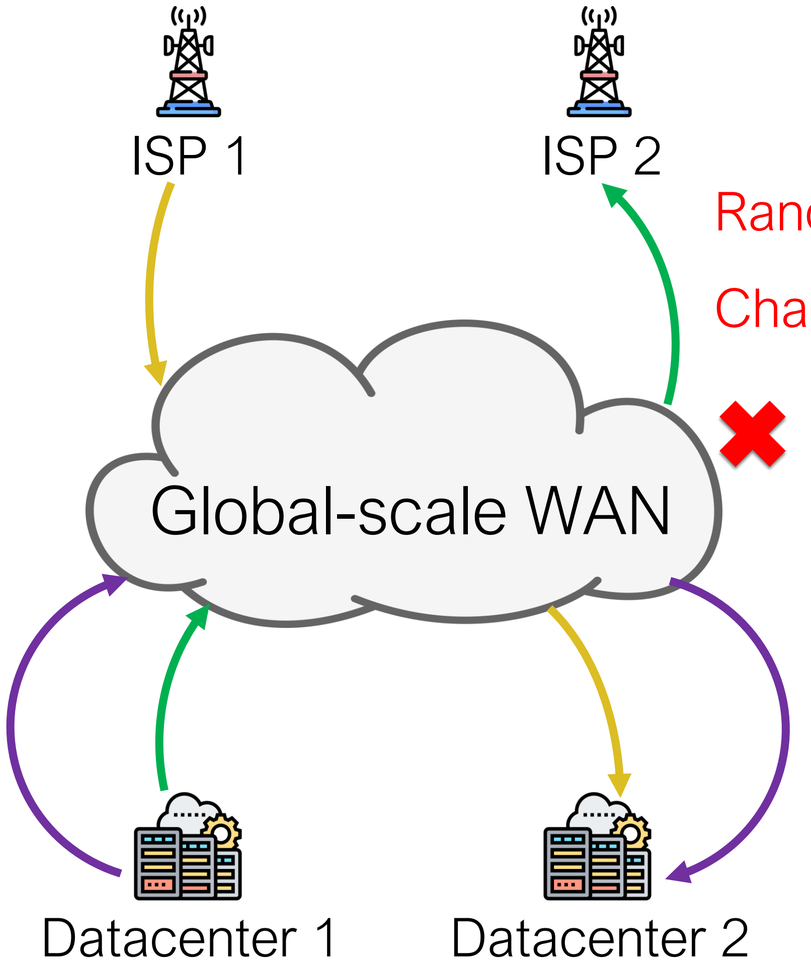
Reasoning about Network Traffic Load Property at Production Scale

Ruihan Li, Fangdan Ye, Yifei Yuan, Ruizhen Yang, Bingchuan Tian, Tianchen Guo, Hao Wu, Xiaobo Zhu, Zhongyu Guan, Qing Ma, Xianlong Zeng, Chenren Xu, Dennis Cai, Ennan Zhai

Overview of our WAN



Random failures

Change plan errors

Cloud computing

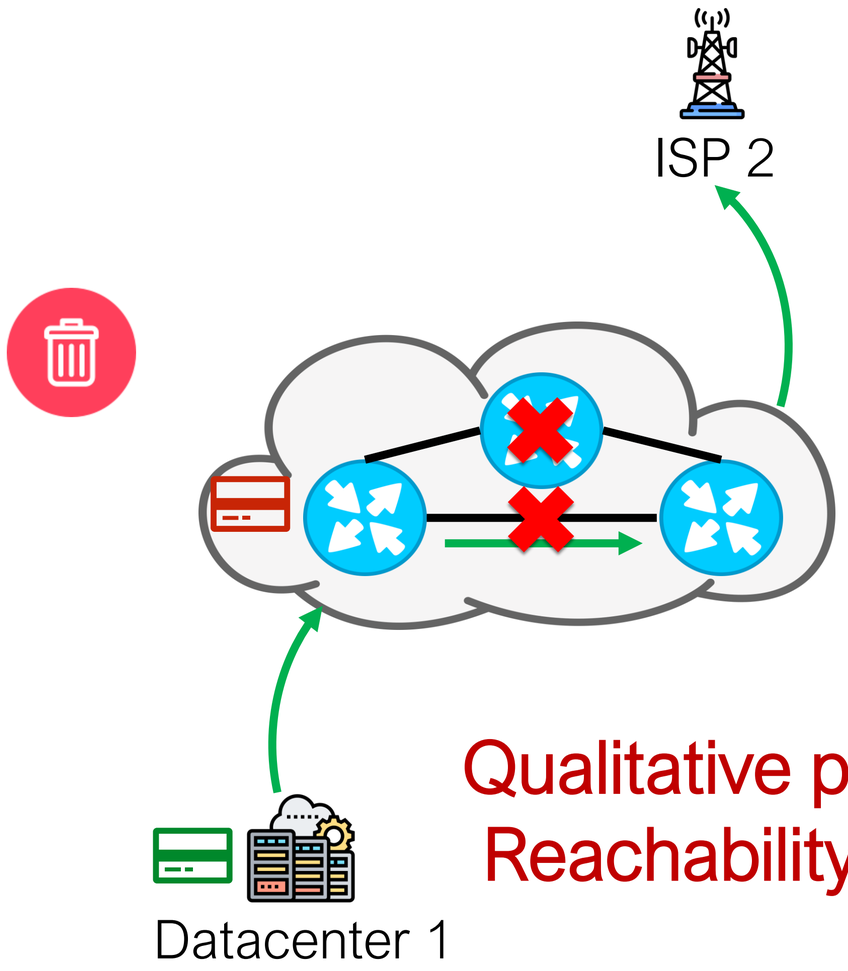
Customer complaints

E-commerce

Economic losses

Proactive outage prevention to the rescue

Background: The state of the art



Qualitative property:
Reachability issues

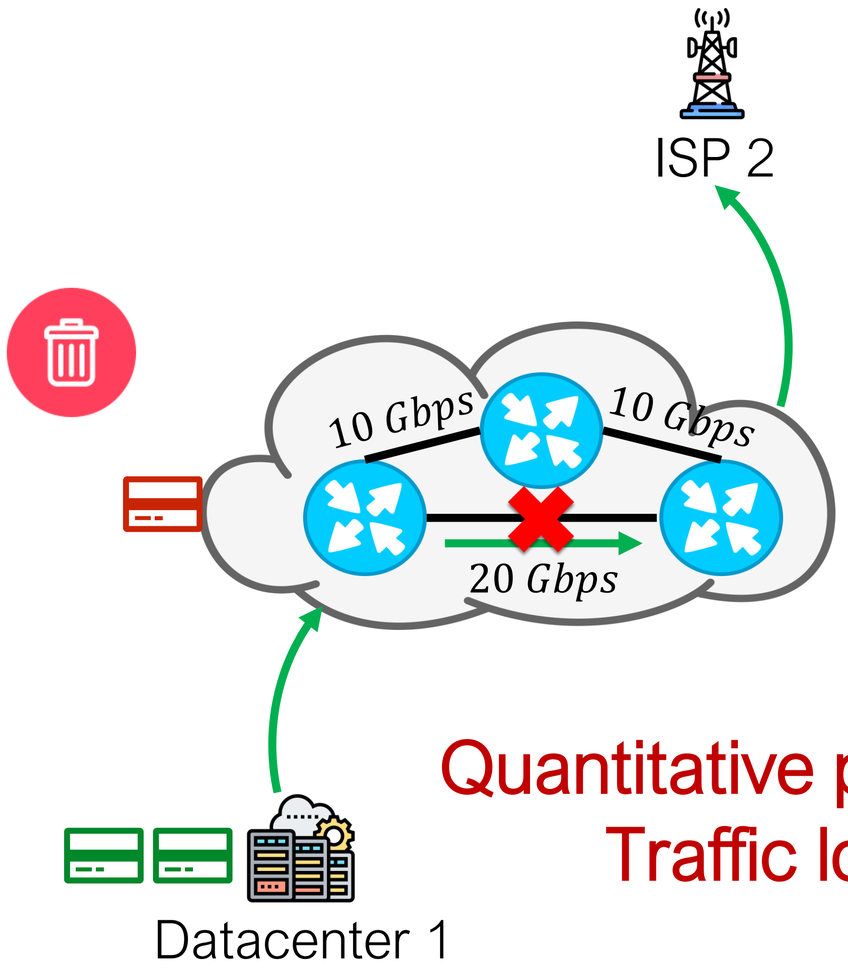
Deployed in our WAN for >3 years
Prevented most reachability errors

Q: What about the remaining?

Batfish [NSDI'15] Hoyan [SIGCOMM'20] ……

Minesweeper [SIGCOMM'17] DNA [NSDI'22] ……

Background: The state of the art



90% of our outages caused by misconfigurations are related to traffic load violations

Jingubang (This work):

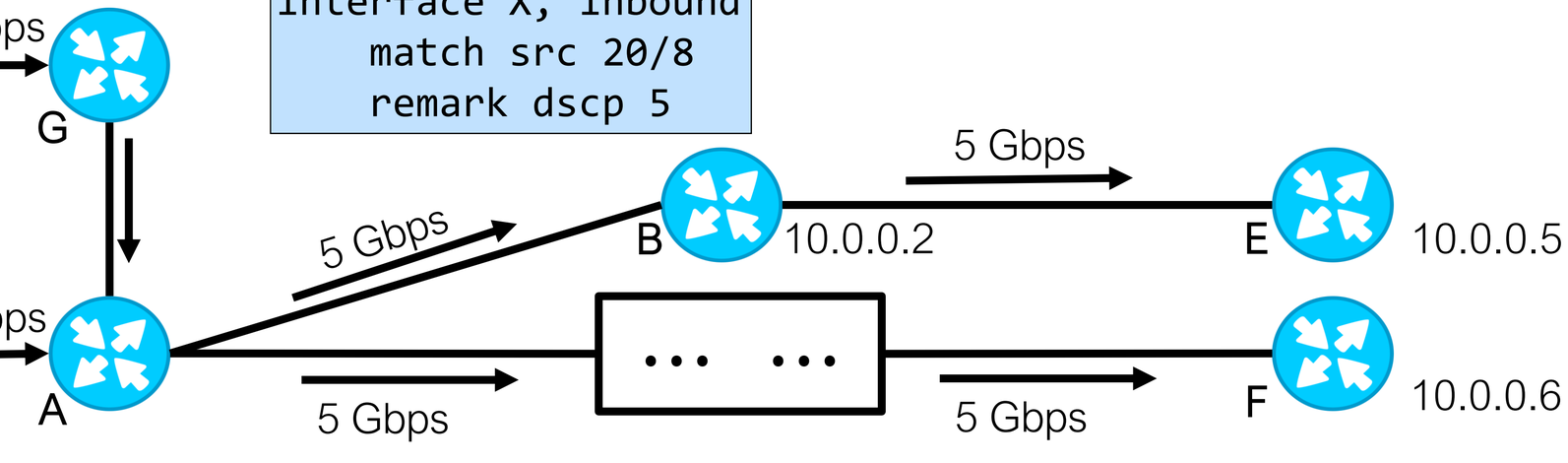
The first system that is able to check traffic load properties at production scale

Quantitative property:
Traffic load

Motivating example

f_2:
$20.0.0.1 \rightarrow 30.0.0.1$
$dscp = 0$

f_1:
$20.0.0.1 \rightarrow 30.0.0.1$
$dscp = 5$

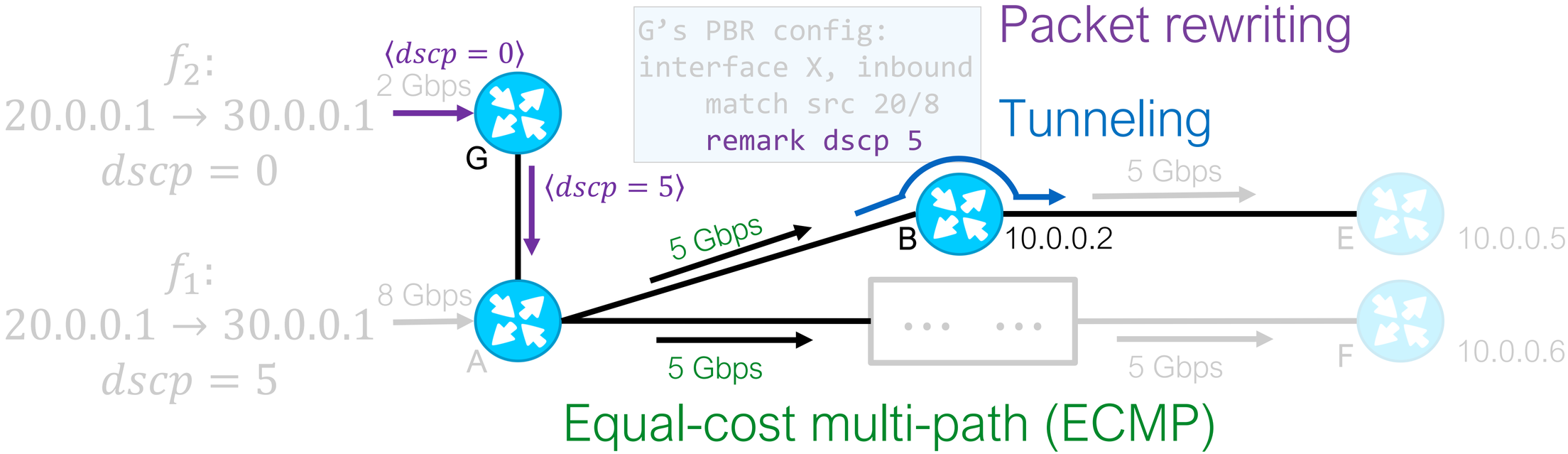



```
G's PBR config:
interface X, inbound
    match src 20/8
    remark dscp 5
```

```
A's BGP RIB:
  *30/8, nexthop 10.0.0.5
  *30/8, nexthop 10.0.0.6
A's IS-IS RIB:
  *10.0.0.5/32, SR tunneled
```

```
A's SR config:
route 10.0.0.5/32, match dscp 5
    path 10.0.0.2, 10.0.0.5 weight=10
```

Challenge 1: Generalizability




```
G's PBR config:
interface X, inbound
    match src 20/8
    remark dscp 5
```

Packet rewriting

Tunneling

Equal-cost multi-path (ECMP)

Distributed control plane

```
A's BGP RIB:
  *30/8, nexthop 10.0.0.5
  *30/8, nexthop 10.0.0.6
A's IS-IS RIB:
  *10.0.0.5/32, SR tunneled
```

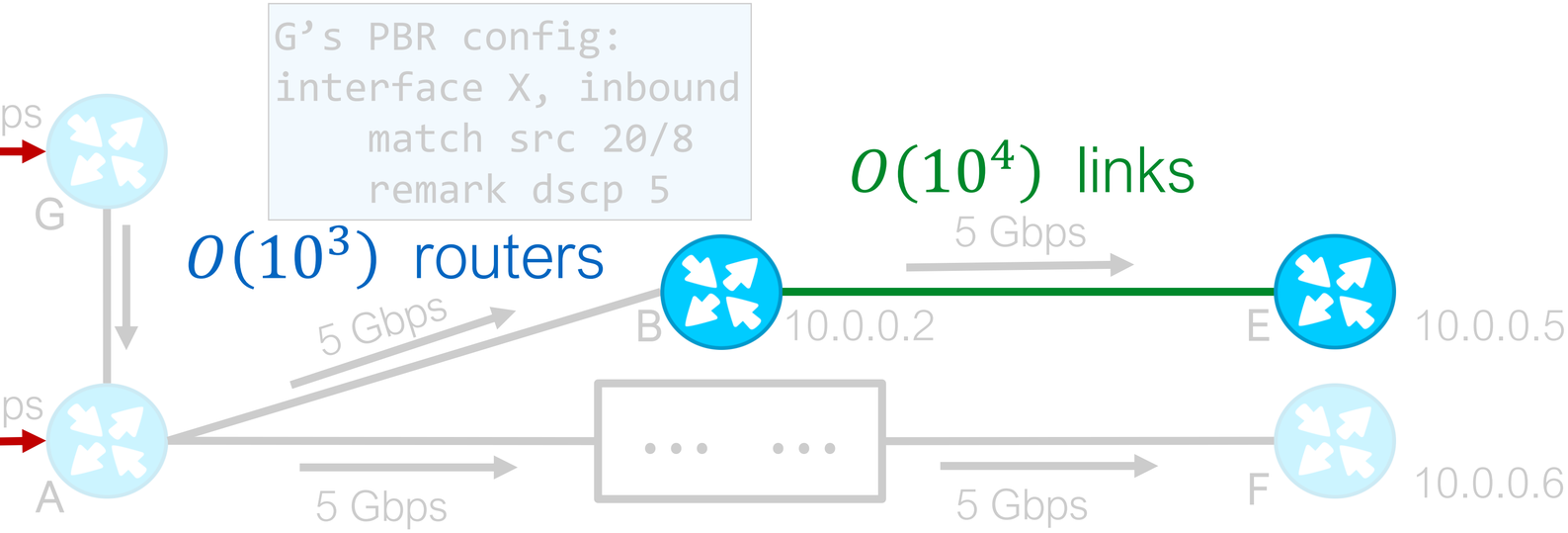
```
A's SR config:
route 10.0.0.5/32, match dscp 5
    path 10.0.0.2, 10.0.0.5 weight=10
```

Simulation ⇒ Comprehensive protocol support

Challenge 2: Scalability

f_2:
$20.0.0.1 \rightarrow 30.0.0.1$
$dscp = 0$

f_1:
$20.0.0.1 \rightarrow 30.0.0.1$
$dscp = 5$



$O(10^9)$ flows in each hour

```
A's BGP RIB:
  *30/8, nexthop 10.0.0.5
  *30/8, nexthop 10.0.0.6
A's IS-IS RIB:
  *10.0.0.5/32, SR tunneled
```

$O(10^6)$ routes on each router

```
A's SR config:
route 10.0.0.5/32, match dscp 5
    path 10.0.0.2, 10.0.0.5 weight=10
```

Novel design & optimization ⇒ WAN usability

Challenge 3: Efficiency

Jingubang is used in: **Multi-stage simulation ⇒ Real-time analysis**

- **Network change analysis:** $O(100)$ changes per week

(full simulation)

- **Failure tolerance analysis:** $O(10^4)$ potential failure cases

(incremental simulation)

Part 1: Network change analysis

Alibaba Cloud
Worldwide Cloud Services Partner

Jingubang is used in: Multi-stage simulation ⇒ Real-time analysis

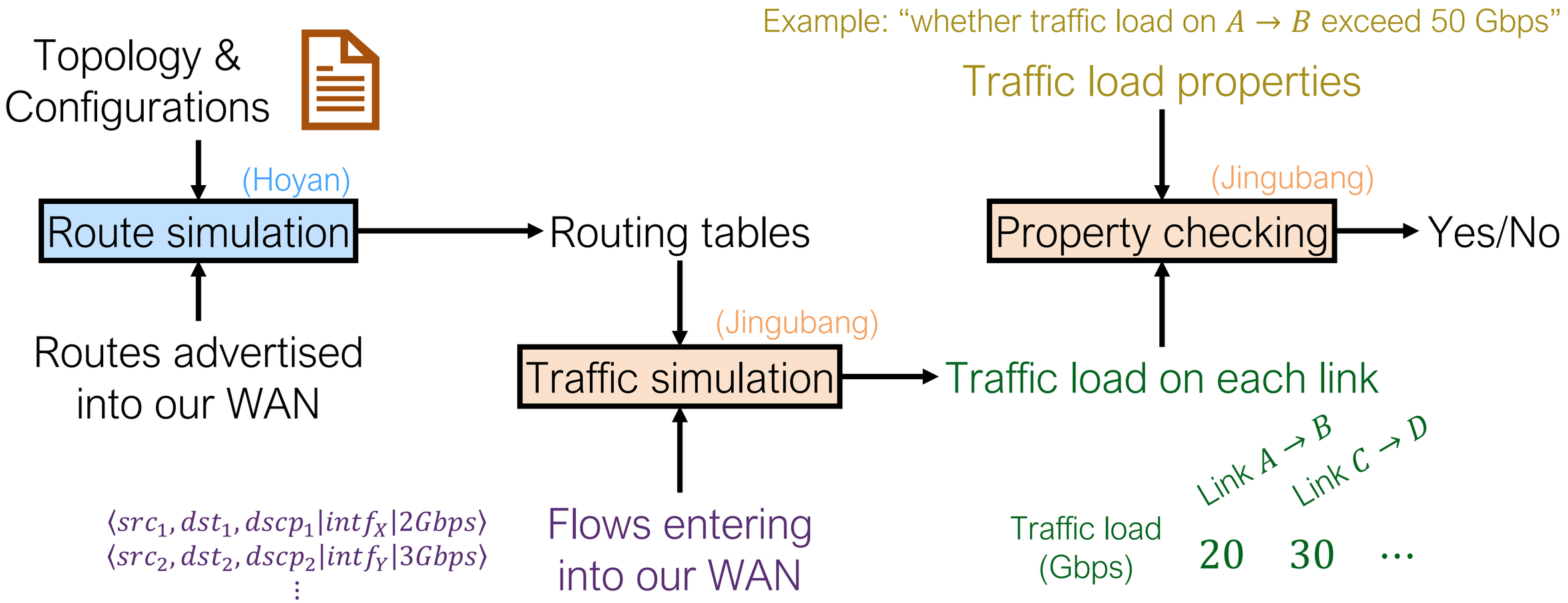
- **Network change analysis:** $O(100)$ changes per week

(full simulation)

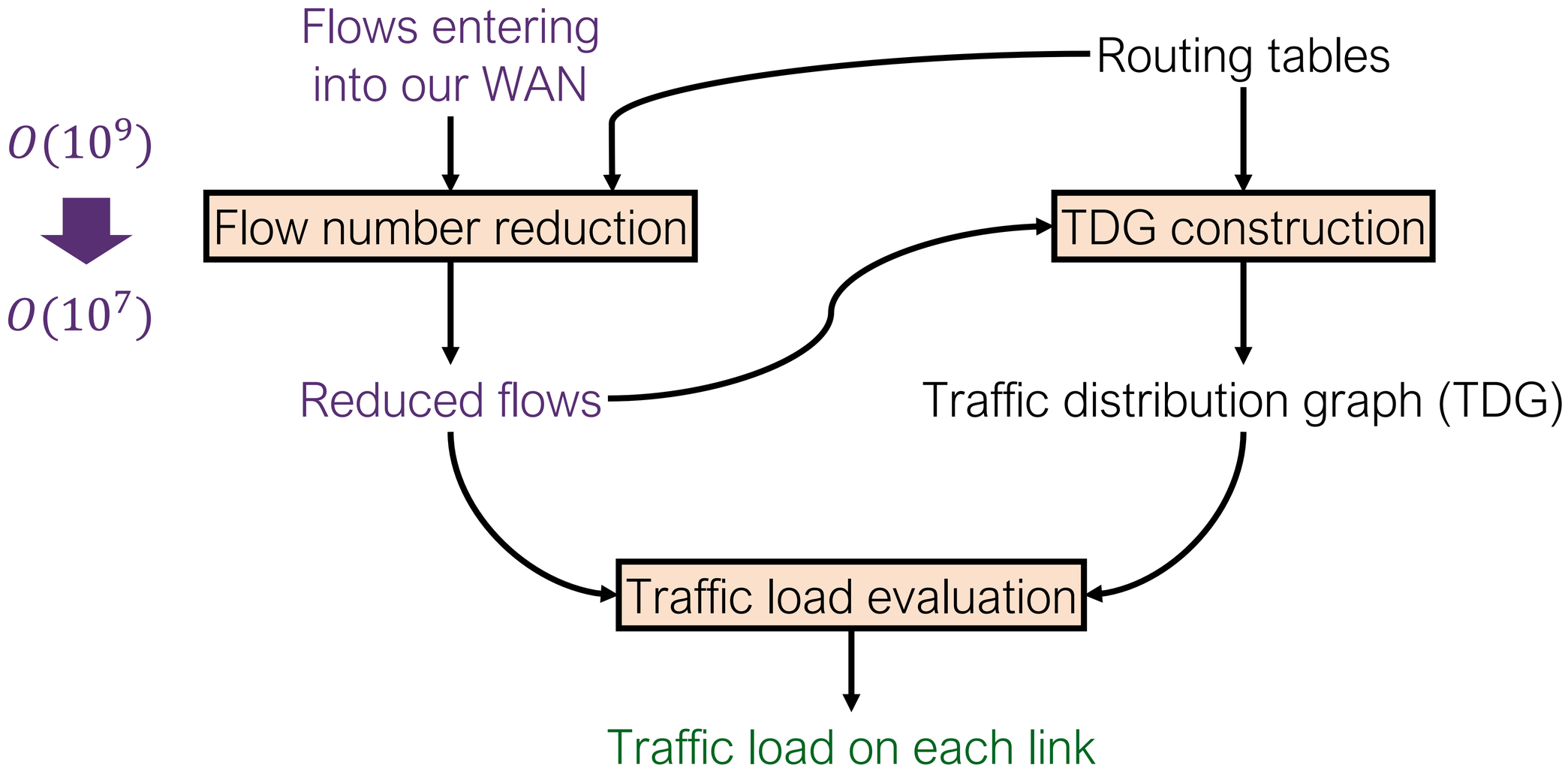
- **Failure tolerance analysis:** $O(10^4)$ potential failure cases

(incremental simulation)

Overview of Jingubang

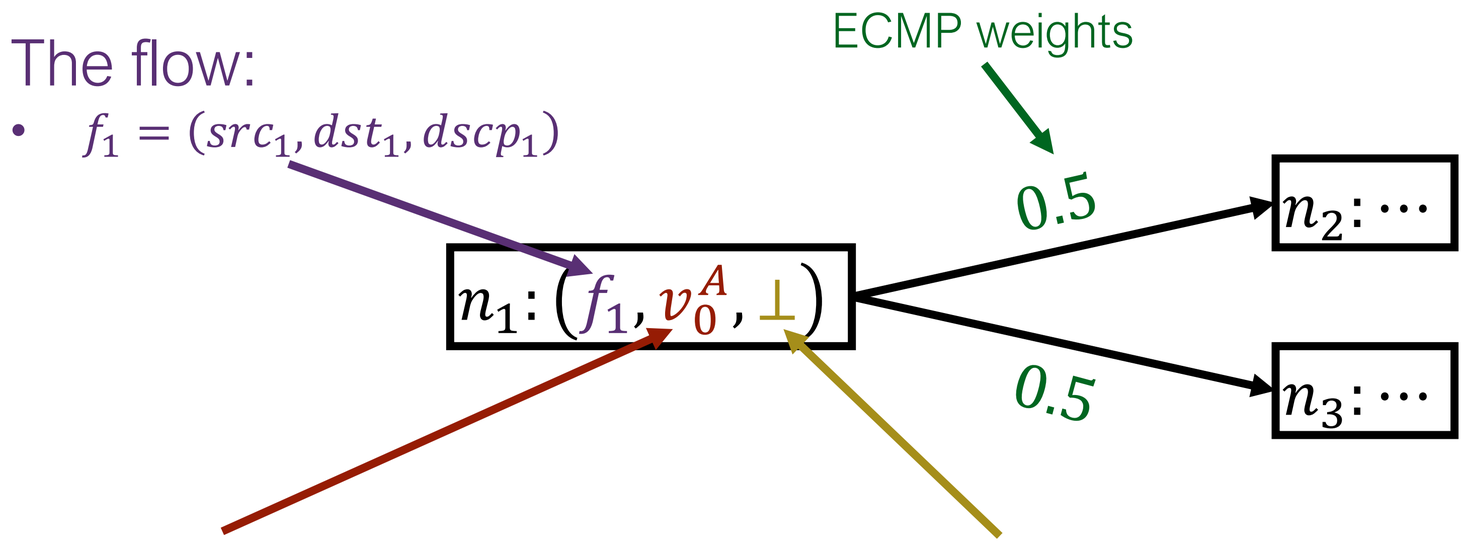


Overview of traffic simulation



TDG Definition

A traffic distribution graph (TDG) is a directed acyclic graph, where:



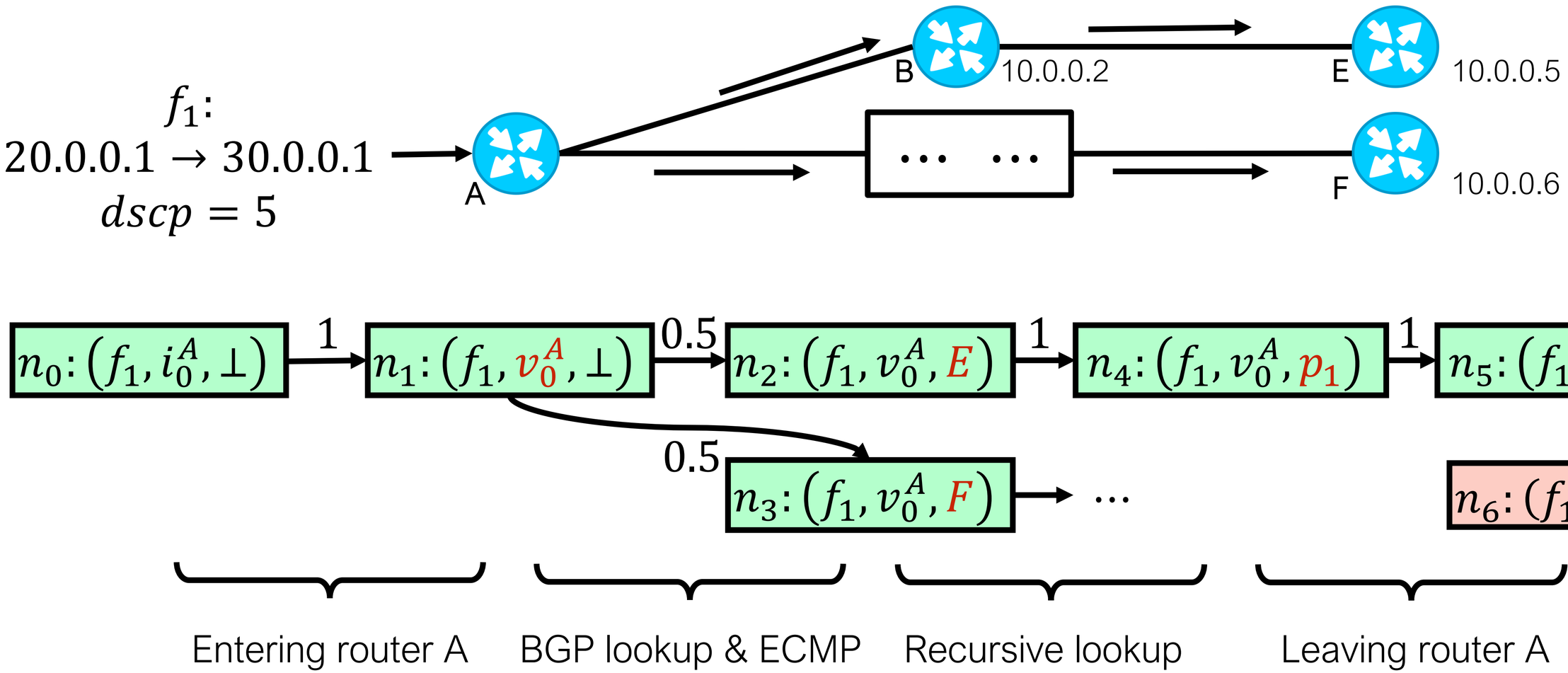
The location of the flow:

- v_0^A: router A
- i_0^A: ingress interface 0
- e_0^A: egress interface 0

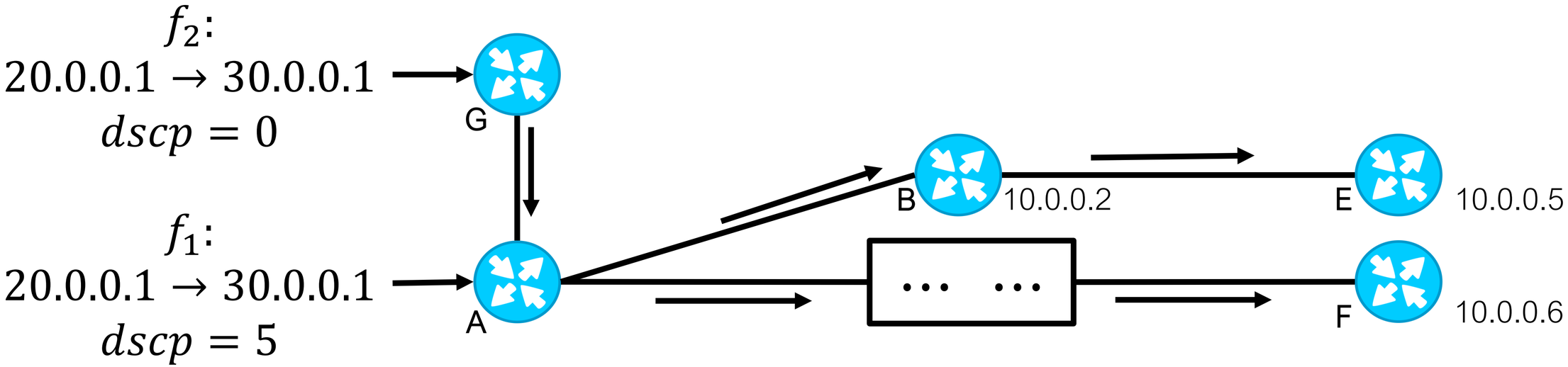
The next-hop of the flow:

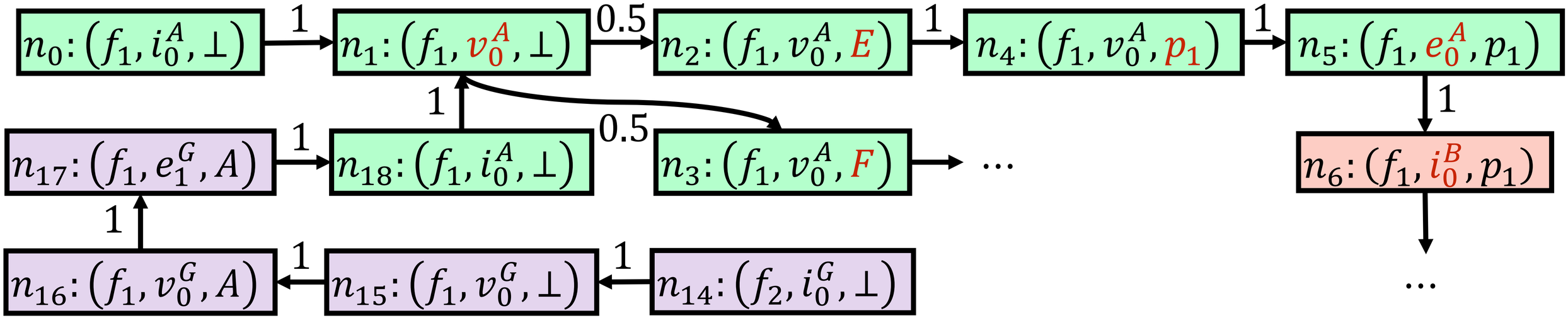
- $\perp$: to be determined
- A: router A
- p_i: tunnel i

TDG construction

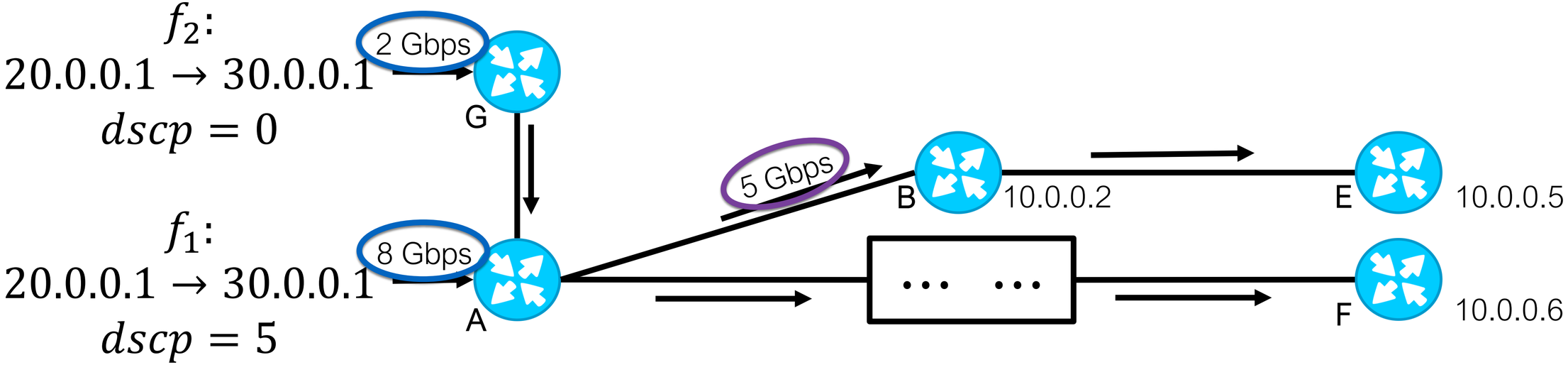


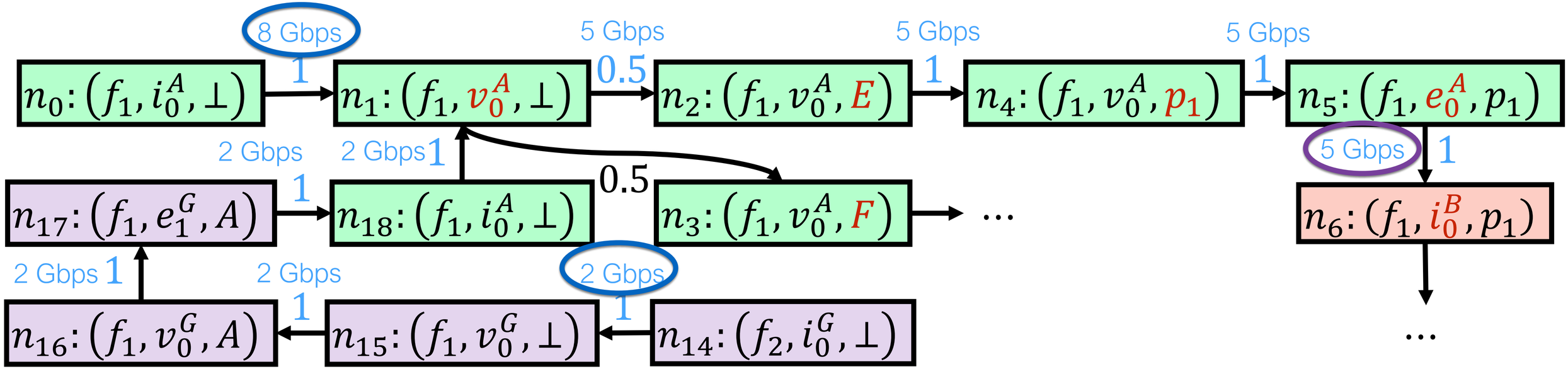
TDG construction



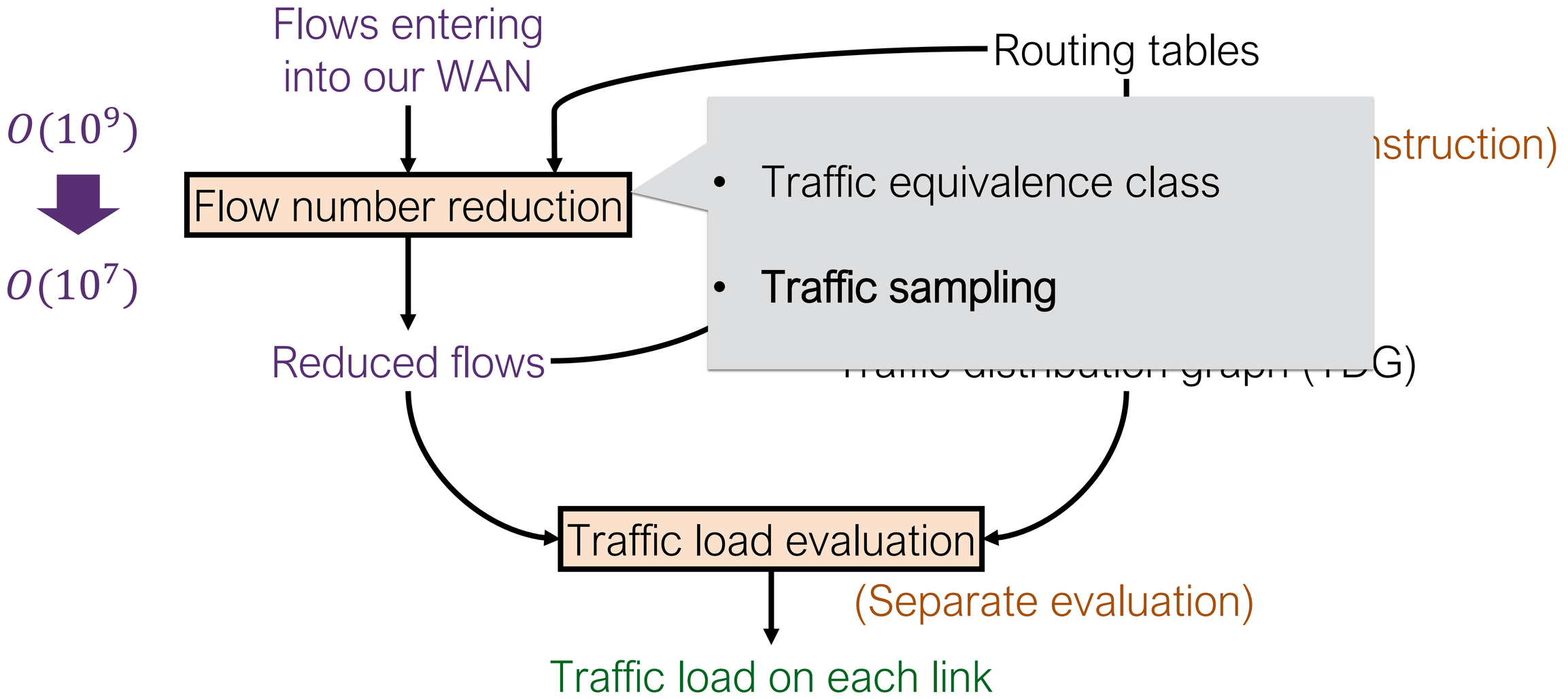


Traffic load evaluation





Overview of traffic simulation



Traffic sampling

Flows entering into our WAN:

$\langle f_1 | int f_X | 8Gbps \rangle$
$\langle f_2 | int f_Y | 2Gbps \rangle$

The total volume of all flows are kept

To sample *at most* 1 flow:

With $\frac{8}{8+2} \times 100\%$ probability, generate $\left\langle f_1 \middle| int f_X \middle| \frac{8+2}{1} Gbps \right\rangle$

With $\frac{2}{8+2} \times 100\%$ probability, generate $\left\langle f_2 \middle| int f_Y \middle| \frac{8+2}{1} Gbps \right\rangle$

} Repeat 1 time

The probability is proportional to the flow's volume

Traffic sampling

Flows entering into our WAN:

$\langle f_1 | int f_X | 8Gbps \rangle$
$\langle f_2 | int f_Y | 2Gbps \rangle$

To sample *at most* k flows:

With $\frac{8}{8+2} \times 100\%$ probability, generate $\left\langle f_1 \middle| int f_X \middle| \frac{8+2}{k} Gbps \right\rangle$
With $\frac{2}{8+2} \times 100\%$ probability, generate $\left\langle f_2 \middle| int f_Y \middle| \frac{8+2}{k} Gbps \right\rangle$

} Repeat k times

Theorem:

Given Δ, μ, δ, we need a k statisfying $k \geq \frac{2V}{\Delta}\left(\frac{1}{\mu} + \frac{1}{3}\right) \ln \frac{2M}{\delta}$

δ: The confidence parameter

Δ: The maximum absolute error

μ: The maximum relative error

V: The total volume
M: The number of links

Part 2: Failure tolerance analysis

Alibaba Cloud
Worldwide Cloud Services Partner

Jingubang is used in: Multi-stage simulation ⇒ Real-time analysis

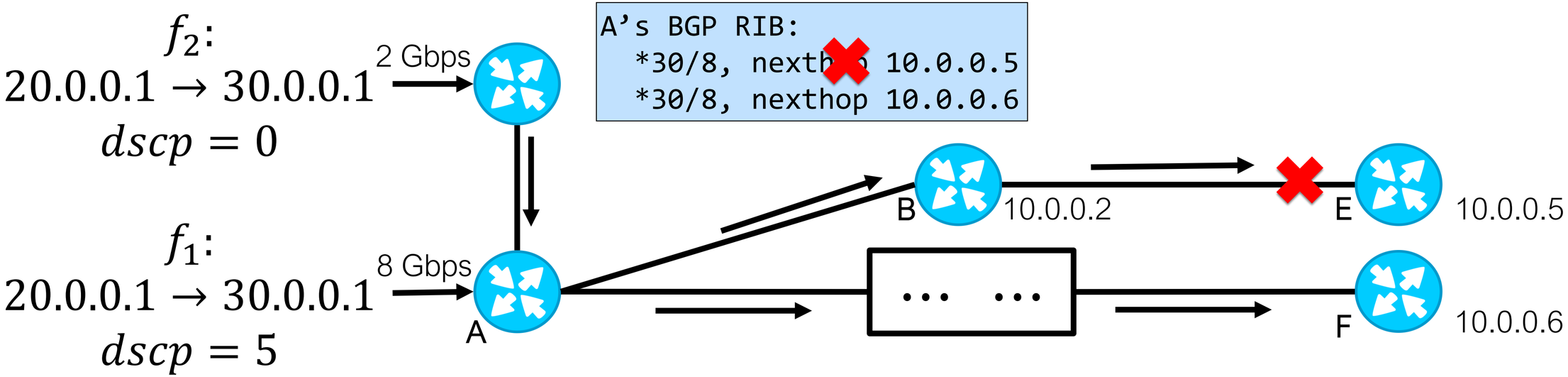
- Network change analysis: $O(100)$ changes per week

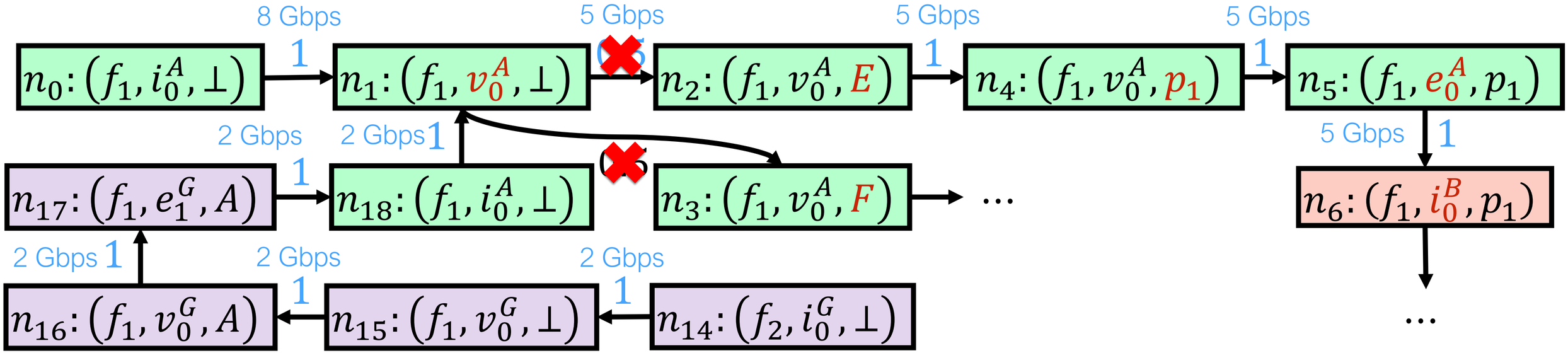
(full simulation)

- **Failure tolerance analysis:** $O(10^4)$ potential failure cases

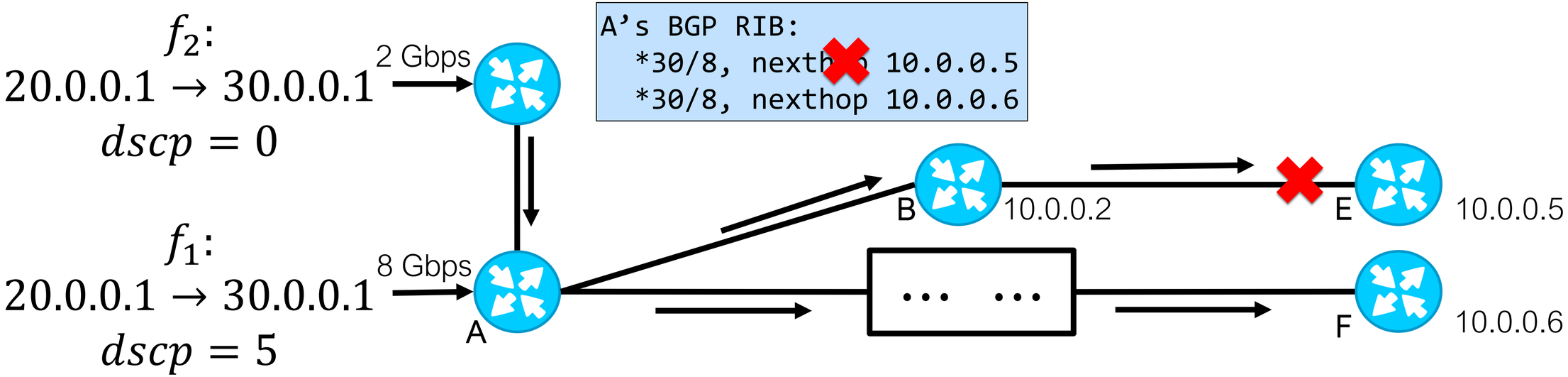
(incremental simulation)

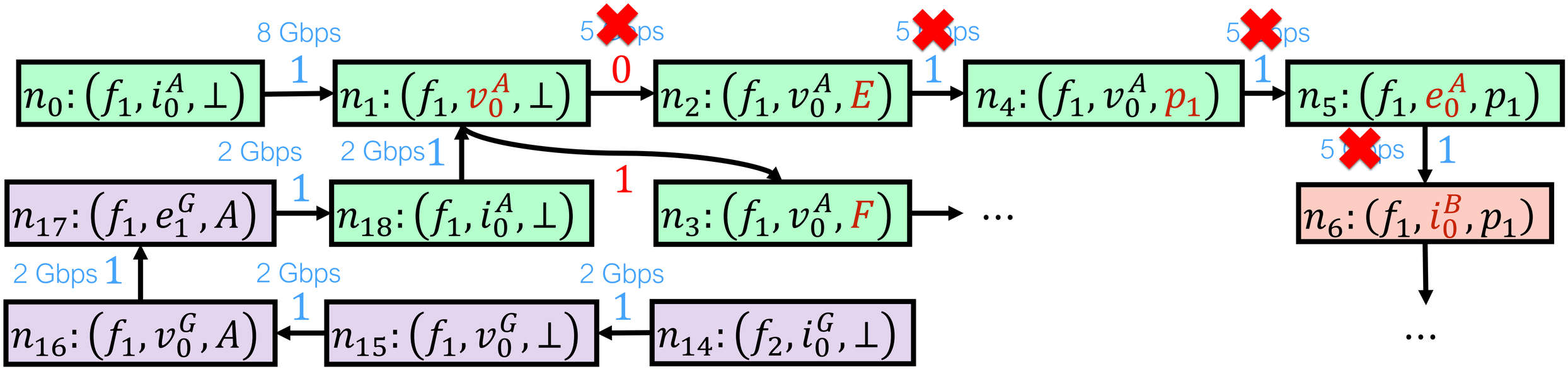
Incremental traffic simulation



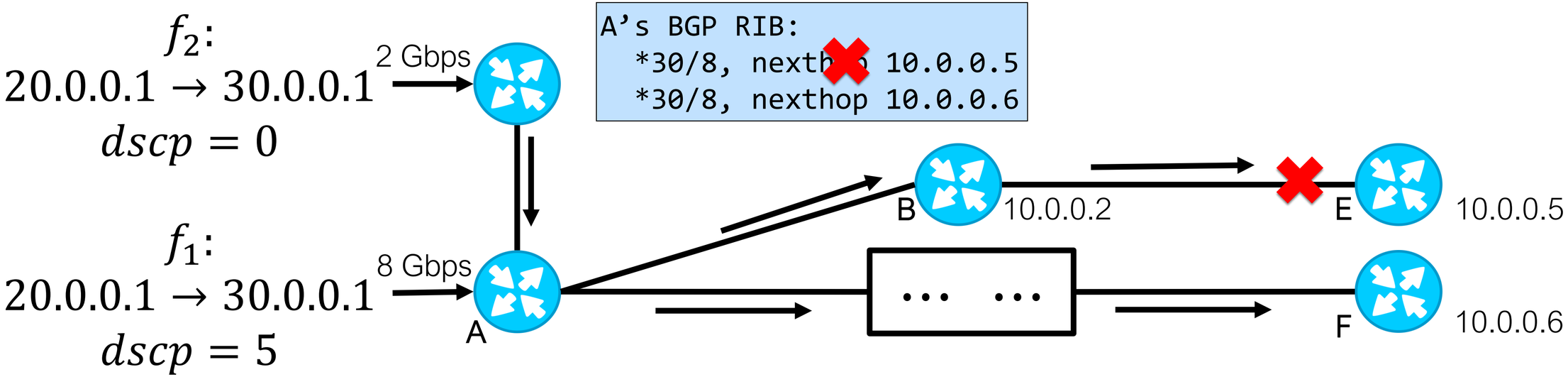


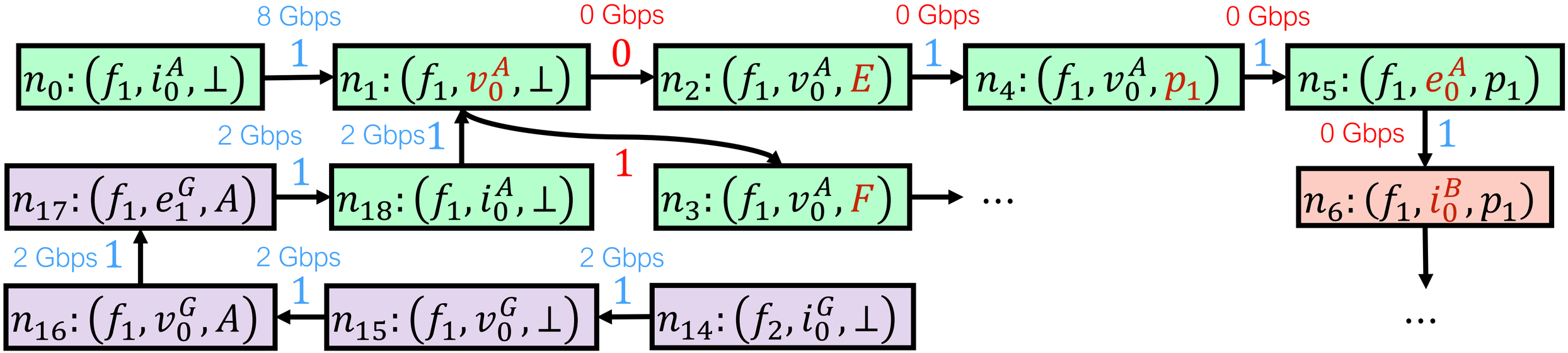
Incremental traffic simulation





Incremental traffic simulation





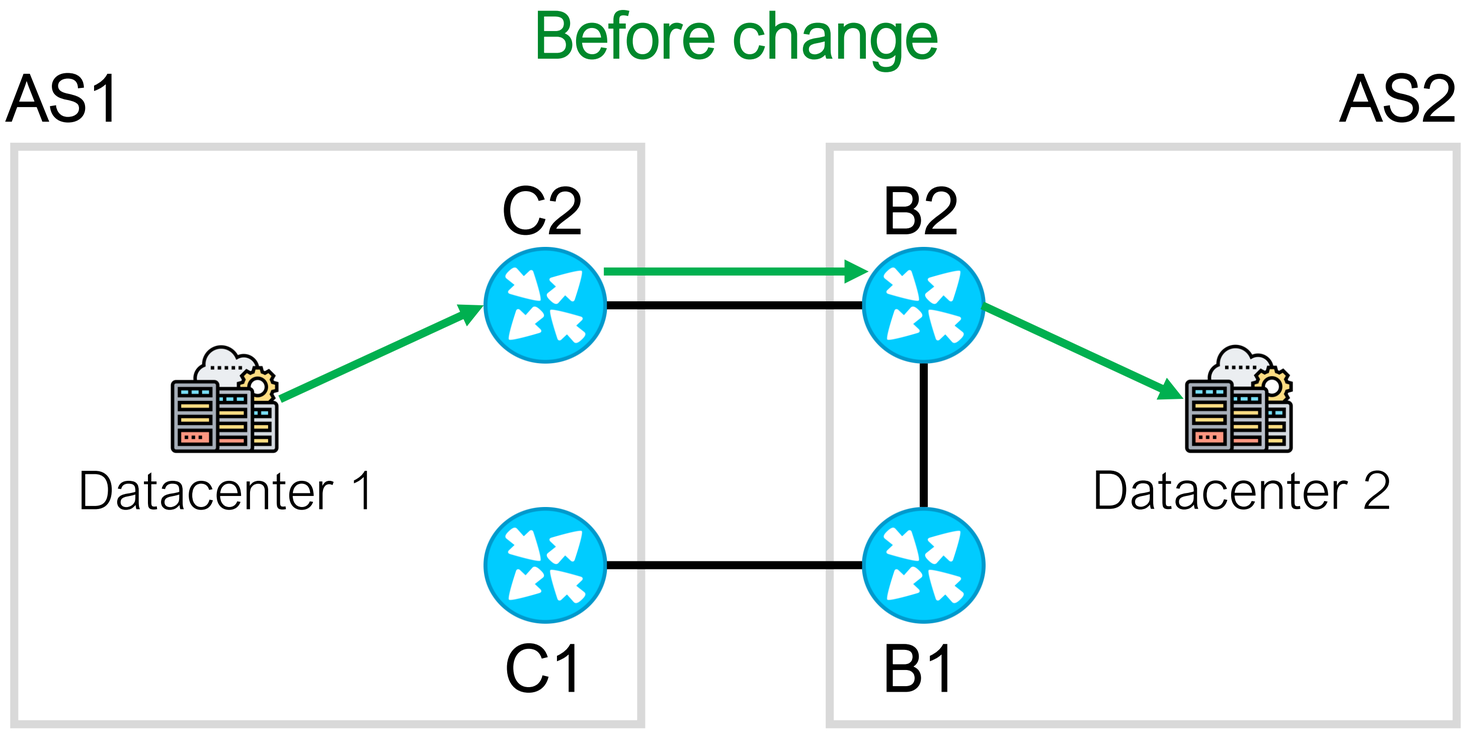
Deployment Experience & Evaluation

Deployment and detected risks

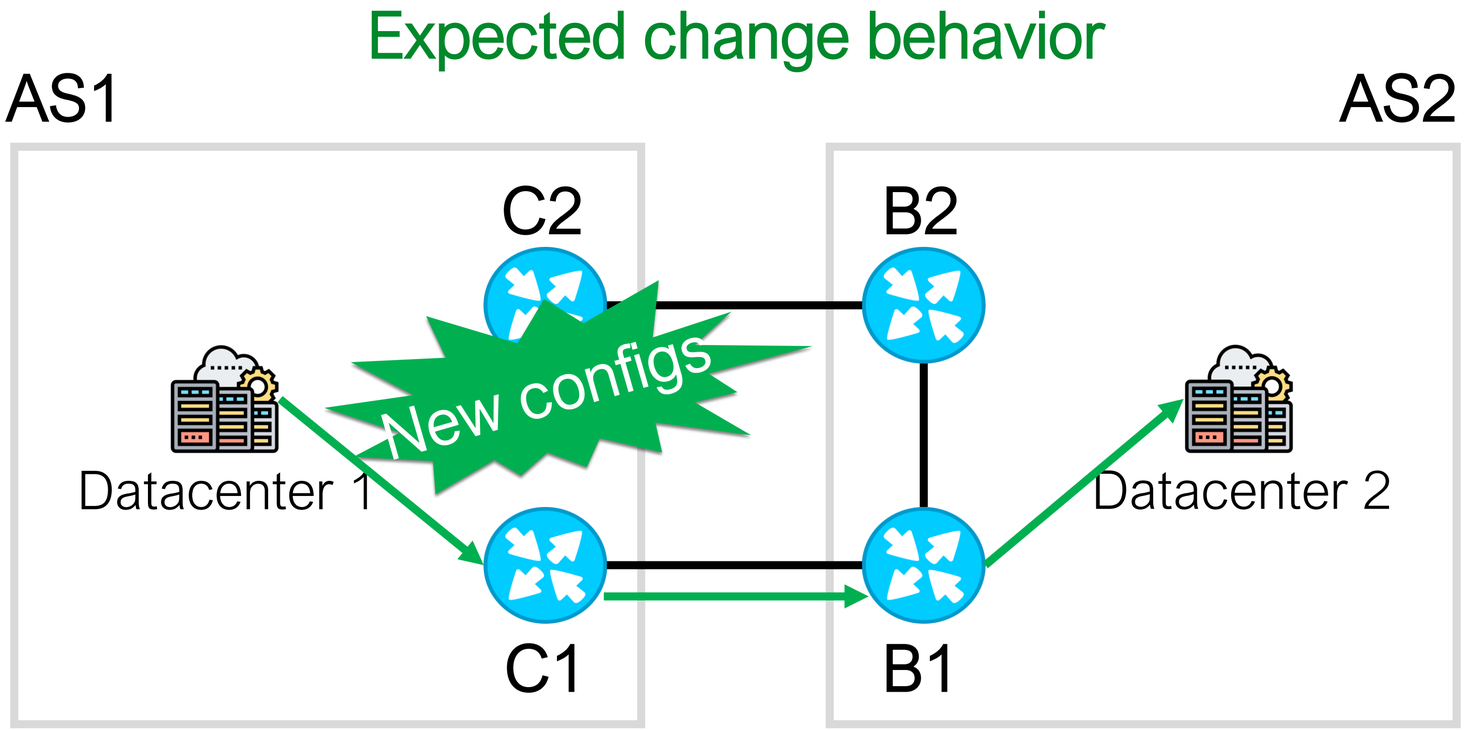
- Jingubang has been deployed in our WAN for more than one year
- Tens of detected risks have been confirmed by our operators

Root causes	Change plan errors	Unexpected routes	Existing misconfiguration
Percentage	44%	33%	23%

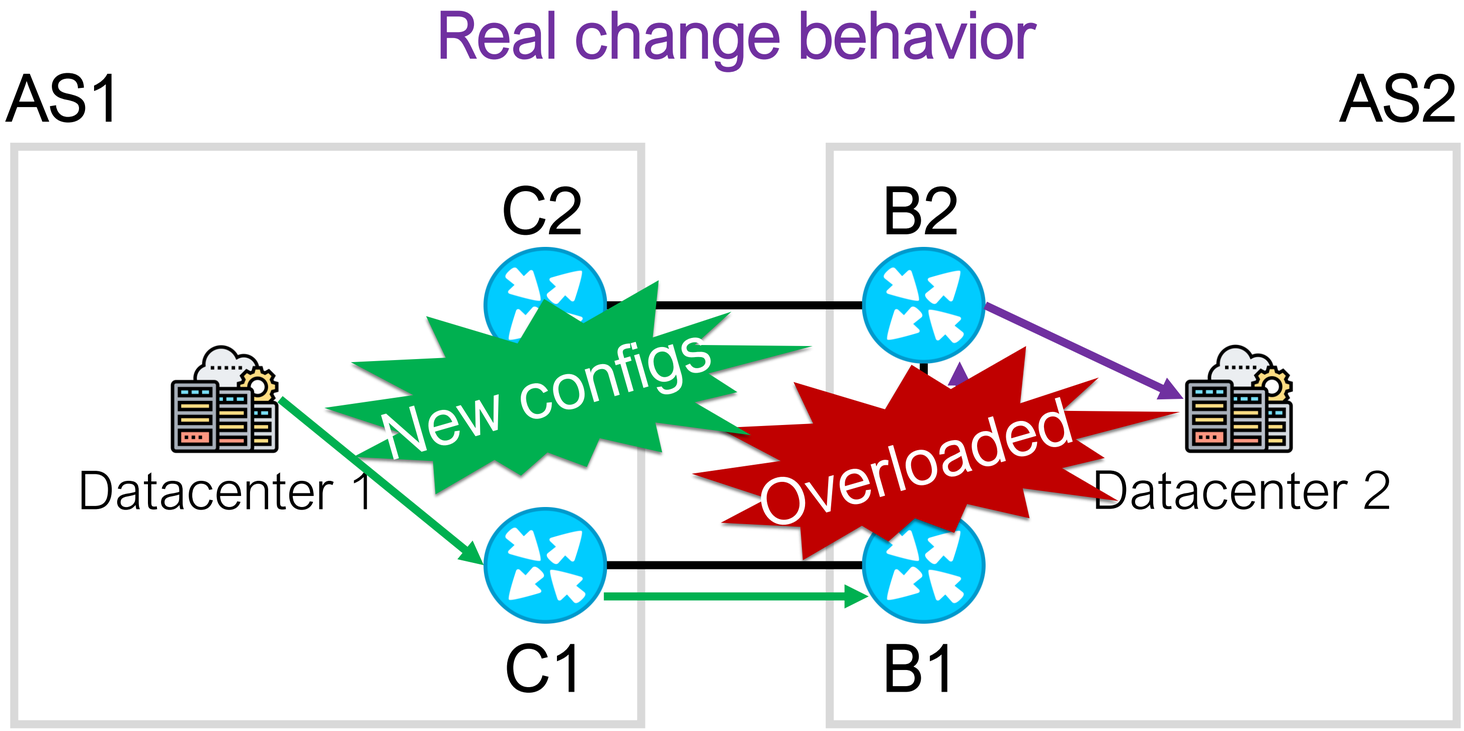
Example: Existing misconfiguration



Example: Existing misconfiguration

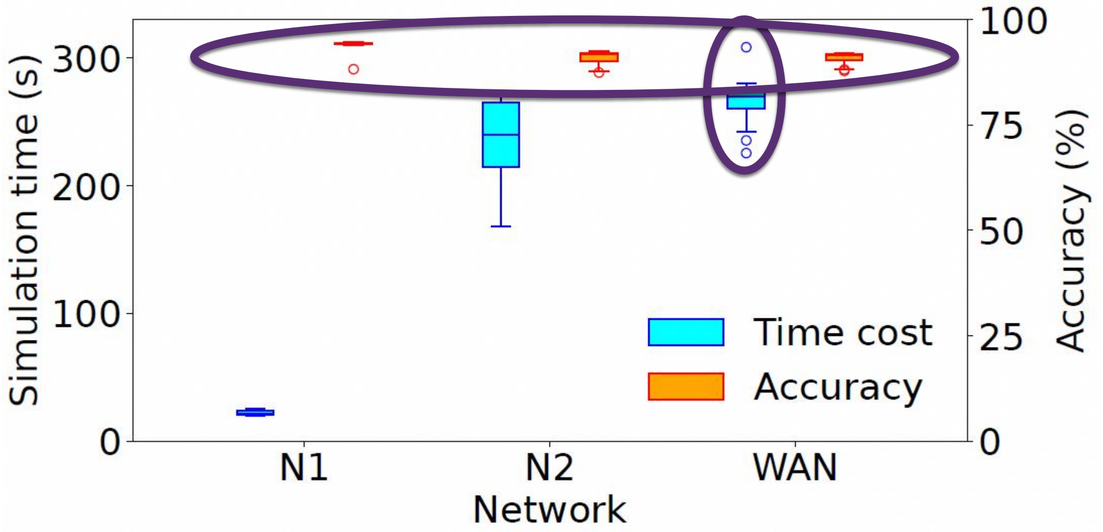


Example: Existing misconfiguration



$$d(B1, DC2) > d(B1, B2) + d(B2, DC1)$$

Evaluation: End-to-end performance & accuracy

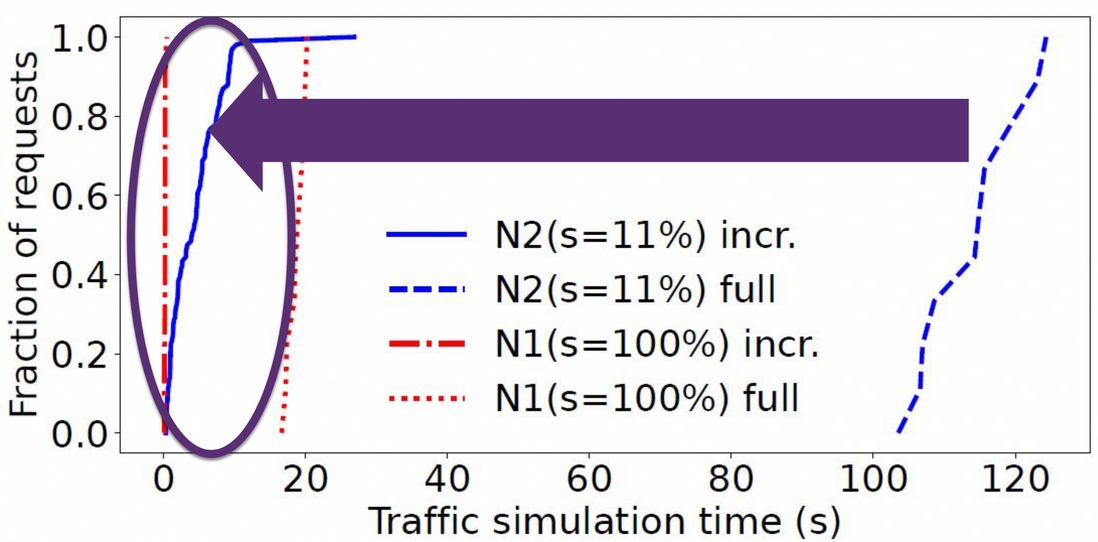


Full simulation costs $\lesssim 5$ min

$\sim 90\%$ accuracy

- $< 100\%$ due to route monitoring

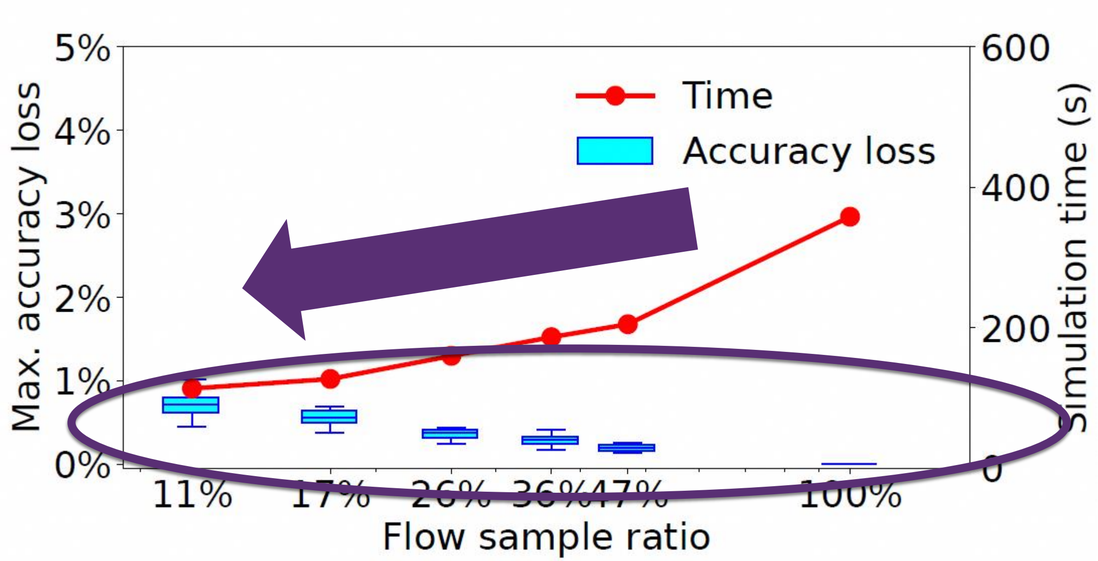
Evaluation: Real-time failure-tolerance analysis



Incremental simulation costs $\lesssim 20$ s

$26.2\times$ speedup

Evaluation: Traffic sampling



3.2 × speedup

Accuracy: For all links, either

- the relative error ≤ 1.1%, or
- the absolute error ≤ 60 Mbps

Conclusion

Significance

- The first system that is able to check traffic load properties at production scale

Results

- We have been using Jingubang in our production WAN for more than one year
- Jingubang has prevented service downtime resulting from traffic load violations

THANKS